DOCUMENT RESUME

ED 107 709 TM 004 536

AUTHOR Brown, Thomas A.
TITLE An Experiment in Probabilistic Forecasting.
INSTITUTION Rand Corp., Santa Monica, Calif.
SPONS AGENCY Advanced Research Projects Agency (DOD), Washington, D.C.
REPORT NO R-944-ARPA
PUB DATE Jul 73
NOTE 52p.

EDRS PRICE MF-$0.76 HC-$3.32 PLUS POSTAGE
DESCRIPTORS College Students; *Current Events; *Decision Making; *Futures (of Society); Measurement Techniques; *Prediction; *Probability; Response Style (Tests); Tests; World Affairs

ABSTRACT

Students were asked to make forecasts of fourteen quantities where true values would not become known for five or six months. The quantities were selected to be typical of the subjects which would be of interest to a decisionmaker in business or government, and included GNP, consumer prices, draft calls, deaths in South Vietnam, and election results on both the state and national level. Rather than giving a one-point estimate for each quantity, each respondent was asked to give a probability distribution which reflected the likely behavior of the quantity in question. It was found that almost all respondents were able to give meaningful distributions and there was a tendency for the true answer to occur disproportionately often in the tails of the distribution given. The effect of this type of error can be counteracted to a certain extent by combining the individual responses into a "consensus" distribution, which will have a greater spread than most of the individuals' distributions. Of the four ways in which this combining of responses was carried out, the most effective was to average the individual's probability density functions. All four consensus methods produced better forecasts than did the average individual. No clear-cut association was found between forescasting skill and age, sex, academic major, or score on a simple current events quiz. (Author/BJG)

ARPA ORDER NO.: 189-1

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE EDUCATION POSITION OR POLICY

TM

R-944-ARPA
July 1973

ED107709

An Experiment in Probabilistic Forecasting

Thomas A. Brown

TM 004 536

A Report prepared for

DEFENSE ADVANCED RESEARCH PROJECTS AGENCY

25th Year
Rand
SANTA MONICA, CA. 90406

The research described in this Report was sponsored by the Defense Advanced Research Projects Agency under contract No. DAHC15-73-C-0181. Reports of The Rand Corporation do not necessarily reflect the opinions or policies of the sponsors of Rand research.

ED107709

ARPA ORDER NO.: 189-1

R-944-ARPA
July 1973

An Experiment in Probabilistic Forecasting

Thomas A. Brown

A Report prepared for

DEFENSE ADVANCED RESEARCH PROJECTS AGENCY

Rand
SANTA MONICA, CA. 90406

4

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

PREFACE

The Rand Corporation is currently making a study of defense issues raised by technological and economic change, under sponsorship by the Defense Advanced Research Projects Agency (ARPA). One of these issues is the question of how the intelligence community can more effectively communicate degrees of uncertainty to decisionmakers. Some researchers at Rand have advocated casting intelligence forecasts into the form of explicit probability statements. Others have contended that individuals have great difficulty making meaningful political or economic forecasts in such terms.

The experiment described in this report was carried out under an ARPA-sponsored Group Judgment Technology project. It is being reported because of its great relevance to the problem of making intelligence forecasts in probabilistic terms. The experiment was designed to test ways of increasing the utility of group and individual forecasts by having them cast in explicitly probabilistic terms. Almost all of the thirty-one respondents (summer students at UCLA) were found to be able to make such forecasts. In a large number of cases the true answer (when it became known) fell in the tail of the distribution forecast by individuals, but this deficiency could be overcome by using a properly selected consensus technique to generate a group forecast.

SUMMARY

On July 28, 1970, thirty-one summer students of the University of California, Los Angeles, were asked to make forecasts of fourteen quantities (where true values would not become known for five or six months). The quantities were selected to be typical of the subjects which would be of interest to a decisionmaker in business or government, and included GNP, consumer prices, draft calls, deaths in South Vietnam, and election results on both the state and national level. Rather than giving a one-point estimate for each quantity, each respondent was asked to give a probability distribution which reflected the likely behavior of the quantity in question.

It was found that almost all respondents were able to give meaningful distributions (95 percent of the responses were usable). There was a tendency for the true answer to occur disproportionately often in the tails of the distributions given (in fact, this occurred twenty times as frequently as one would "expect"). The effect of this type of error can be counteracted to a certain extent by combining the individual responses into a "consensus" distribution, which will have a greater spread than most of the individuals' distributions. Of the four ways in which this combining of responses was carried out, the most effective (and the simplest conceptually) was to average the individual's probability density functions. All four consensus methods produced better forecasts than did the average individual.

No clear-cut association was found between forecasting skill and age, sex, academic major, or score on a simple current events quiz.

CONTENTS

PREFACE iii

SUMMARY v

Section
I. INTRODUCTION 1

II. THE QUESTIONS 3

III. THE RESPONDENTS 5

IV. THE RESPONSES 7

V. THE SCORING SYSTEM 11

VI. CONSENSUS FORECASTS 14

Appendix
A. MAIN QUESTIONNAIRE 19
B. CURRENT EVENTS BACKGROUND QUIZ 36

REFERENCES 45

I. INTRODUCTION

With the rise of decision theory and the increasingly quantitative approach to problems of business and government, there is continuing interest in determining not only what is *most likely* to happen in the future, but just *how likely* various alternative contingencies may be. Of course, throughout history wise decisionmakers have taken a full spectrum of alternative contingencies into account in laying their plans. More than two millenia ago Herodotus wrote: "The best man in my belief is he who lays his plans warily with an eye for every disaster which might occur, and then when the time for action comes, acts boldly."

However, the acceptance of explicit probability distributions over a broad set of contingencies as a useful way to think about real-world problems is comparatively recent. Many intelligent people still question the utility of such specifically quantitative approaches to the kind of uncertainties that arise in business, politics, and war. They argue that it makes no epistemological sense to estimate the "probability" of an inherently unique event such as the outbreak of war before a certain date or the victory of a given candidate in a coming election. In addition, they point out that selecting the contingencies to be considered is a more difficult and creative task for the forecaster than simply assigning probabilities to the elements of some given spectrum.

Finally, they argue that the forecaster should deepen the decisionmaker's *understanding* of the problem, and that this can only be done by conveying to him the structure and underlying dynamics of the situation he faces. If the decisionmaker were simply playing roulette, all he would have to know about 00 was that it had one chance in thirty-eight of coming up; but an election, for example, is not a roulette wheel but a complex interaction of a variety of factors, about each of which the decisionmaker may be constantly receiving information. If you tell him that Abu Ben Said has one chance in thirty-eight of winning a given election, you have not provided nearly as valuable a service as you have if you explain to him how the various forces in the country in

question interact in order to reduce Mr. Said's chances to a rather low level.

The epistemological question of what is "meant" by the probability of an inherently unique event is a difficult one. But in real life we make bets on horse races, football games, and elections (each of which is inherently unique); weather forecasters, aided by the theory of reproducing scoring systems,[1,2] speak of the probability of rain on a certain day; and from a common-sense standpoint it seems unreasonable to be inhibited from using the language of probability theory by purely philosophical considerations.

Those who argue against probabilistic forecasts on the grounds that a different product would be more useful to the decisionmaker are right in some cases, but sometimes they fail to appreciate the decisionmaker's point of view. For example, a corporate treasurer may want to know whether interest rates will rise or fall, but have little need to know what the reasons behind the fluctuations might be. If you are deciding whether to go on a picnic tomorrow, a precipitation forecast may be of interest to you, but an hour's lecture on the principles of meteorology might be a waste of your time.

Even if you are convinced of the meaningfulness and value of explicitly probabilistic forecasts, there are still a number of technical problems to be solved.

- How should you frame the questions you put to your advisors?
- Are most people capable of framing forecasts in probabilistic terms?
- Are there any general biases that individuals exhibit when asked to make such forecasts?
- How should forecasts on the same topic made by different individuals be combined into a "consensus" forecast?

These are empirical questions, and the purpose of this report is to describe an experiment carried out at The Rand Corporation which sheds some light on these questions.

II. THE QUESTIONS

Numerous experiments have been conducted asking respondents to make bets on various random events involving odd-shaped dice, special roulette wheels, and so on.[3,4] However, we wanted to make the questions we asked our subjects be as close as possible to the sort of questions a "real-life" decisionmaker in business or government would be likely to ask.

We selected fourteen quantities which reflect economic, political, and military trends, and asked the respondents for (roughly) six-month forecasts. The experiment was carried out on July 28, 1970. The specific questions were as follows:

1. What will be the value of Standard and Poor's 500-stock average in December 1970?
2. What will be the annual value of the U.S. Gross National Product during the fourth quarter, 1970?
3. What will be the annual rate of the U.S. Gross National Product in 1958 dollars during the fourth quarter of 1970?
4. Where will the consumer price index stand in December 1970?
5. What will be the unemployment rate in December 1970?
6. What will be the annual rate of U.S. military spending in the fourth quarter of 1970?
7. What will be the total Selective Service call for December 1970?
8. How many U.S. military deaths will occur in Southeast Asia from hostile causes in December 1970?
9. How many U.S. military personnel will be in Vietnam in December 1970?
10. How many Republican governors will there be in the United States after the election of November 1970?
11. What percentage of the popular vote will the Republican candidate for U.S. Senator from California receive in the November 1970 election?

12. What percentage of the popular vote will the Republican candidate for Governor of California receive in the election of November 1970?
13. How many Democrats will there be in the House of Representatives after the election of November 1970?
14. How many Democrats will there be in the Senate after the election of 1970?

The respondents were provided with considerable background information on each of these questions. The response pattern desired was the same for each question:

1% chance of being less than _____.
10% " " " " " _____.
30% " " " " " _____.
50% " " " " " _____.
70% " " " " " _____.
90% " " " " " _____.
99% " " " " " _____.

The questionnaire administered to the respondents is reproduced in Appendix A. "Question 0" was a sample question, part of the instructions intended to explain the type of response desired. The final question, on the outcome of the Paris peace talks, has been omitted from the analysis because the response format was different from that in the other fourteen questions.

III. THE RESPONDENTS

The thirty-one respondents were recruited through advertising in the UCLA campus newspaper. They were all summer students at UCLA and were paid a flat fee for participating in the experiment. The majority (16 of 31) were graduate students; there were 14 upper classmen and one freshman. They ranged in age from 19 to 44; their majors were varied.

The respondents were divided into two groups (a morning and an afternoon group) because of space limitations. The characteristics of the various respondents are summarized by group in Table 1.

A quiz on current events, designed to test the respondents' familiarity with the subjects covered in the questionnaire, was administered. This quiz consisted of 45 multiple-choice questions, and the scores achieved ranged from 13 (a freshman math major) to 37 (a senior in history). The quiz is reproduced in its entirety in Appendix B.

It is now our opinion that some questions were ambiguous, and some were so recondite that they shed no light on the relative knowledgeability of the respondents; nevertheless, the quiz does provide a rough and ready measure of their familiarity with the subjects treated in the forecasting questions. By the way, so much information was given to the respondents with each question that a clever respondent could conceivably make moderately good forecasts without knowing very much a priori about the subject in question.

Table 1

CHARACTERISTICS OF RESPONDENTS

Respondent Number	Sex	Age	Academic Status	Score (on current events quiz)	Major
Morning Group					
1	F	30	Grad.	18	Counselor Ed.
2	F	41	Grad.	29	Counselor Ed.
3	F	19	Fresh.	13	Mathematics
4	F	24	Sr.	19	Sociology
5	F	23	Grad.	15	Opera
6	F	21	Sr.	20	Sociology
7	F	23	Grad.	18	Psychology
8	F	21	Sr.	19	Psychology
9	F	20	Jr.	19	Occup. Therapy
10	M	29	Grad.	28	Spanish
11	M	22	Grad.	33	Computer Sci.
12	F	20	Sr.	27	Zoology
13	F	23	Grad.	20	Special Ed.
14	F	28	Grad.	30	Social Work
15	M	23	Grad.	29	Recreation
Afternoon Group					
1	M	25	Sr.	35	Anthropology
2	M	21	Grad.	24	Psychology
3	M	21	Sr.	37	History
4	M	20	Sr.	36	Zoology
5	M	18	Jr.	34	Political Sci.
6	M	26	Grad.	22	English
7	F	21	Sr.	20	Dance
8	M	22	Sr.	17	Business
9	F	27	Grad.	28	History of Art
10	F	44	Grad.	NA[a]	Public Health
11	M	24	Grad.	24	Social Anthro.
12	M	23	Grad.	19	Sociology
13	F	24	Grad.	22	Sociology
14	M	20	Sr.	26	Economics
15	M	20	Sr.	18	Economics
16	M	24	Sr.	26	Political Sci.

[a]Not available.

IV. THE RESPONSES

On examining our response format, several people experienced in psychological experiments commented that our respondents would find our response format very difficult, and that many responses would be unusable. A number of respondents did seem to find the task extremely difficult, but most of them did not. A breakdown of the responses is as follows, for 14 responses from each of 31 subjects:

	Number	Percent
Usable responses	414	95.4
Unusable responses	20	4.6
Total (14 × 31)	434	100.0

Twenty-five of the thirty-one respondents were able to generate meaningful responses to all fourteen questions. Three of the respondents accounted for fifteen of the unusable responses. We called a response "unusable" if it was partially blank, if it was non-monotone, or if the units were not clearly marked by the respondent.

One striking characteristic of the responses, which has also appeared in other experiments of a similar nature, is that the true answer appeared in the tails of the respondents' distributions much more frequently than the respondents would have expected. This is shown graphically in Fig. 1.

In 114 cases, the true answer fell below a point someone has selected as the 1-percentile level; in 61-1/2 cases (the 1/2 reflects an answer right on the line) the true answer fell at a point greater than a point someone had selected as the 99-percentile level. The respondents implicitly claimed to expect events of each of these types to occur only 4.14 times (414 × .01).

This tendency seems to be common among individuals asked to make probability assessments. For example, an unpublished paper by Alpert and Raiffa[5] reports on their efforts to overcome this deficiency. They had about a thousand respondents, so their data base is much

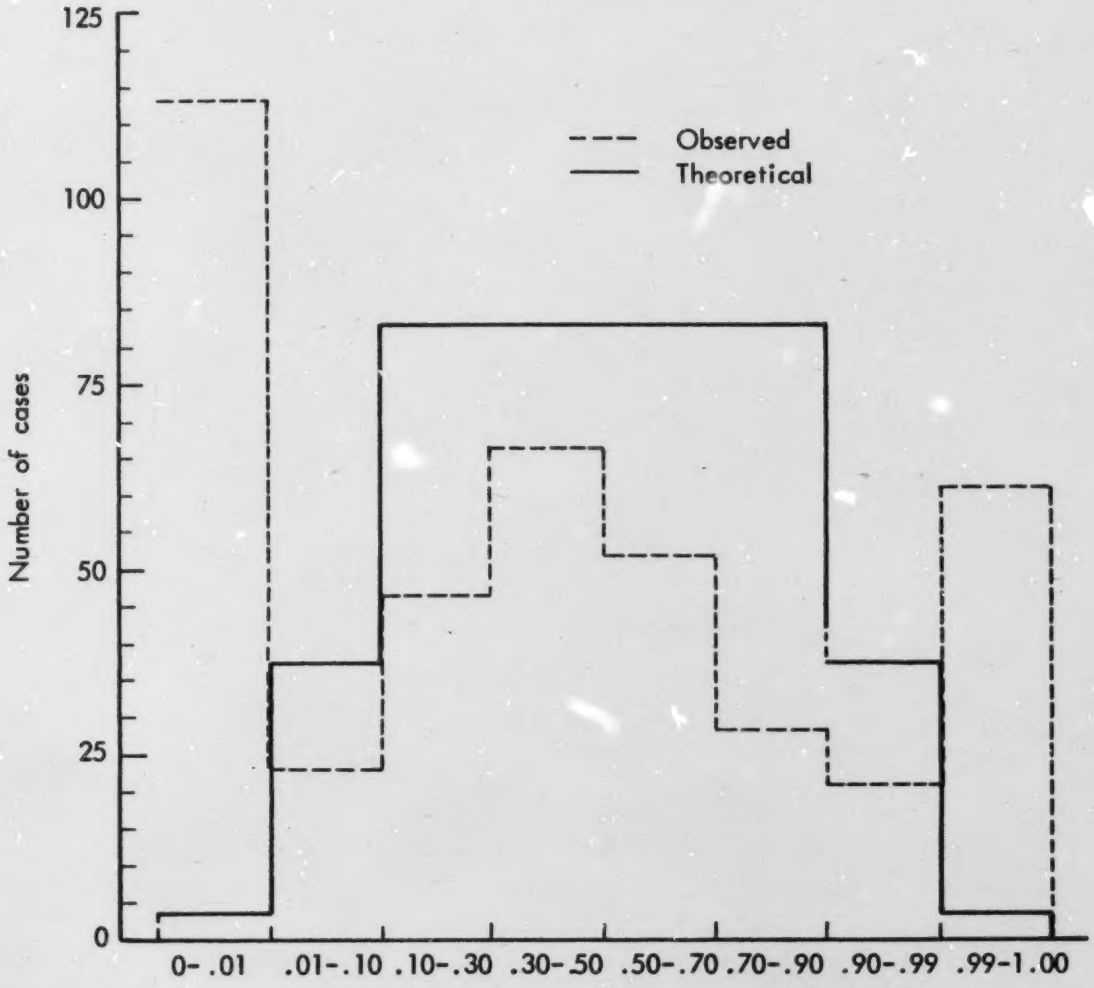


Fig. 1 — Intervals in which true answer falls

larger than ours. They asked the respondents to report their subjective probability distributions for uncertain quantities in a form rather similar to ours. Half of their respondents were warned, in very strong terms, of the tendency to make the .01- and .99-percentile breaks too "tight." Half were not so instructed. Their results compare with ours as follows:

	Percentage of True Values Falling Outside the .01 to .99 Range
Brown (1970)	42
Alpert-Raiffa (1969), uninstructed	41
Alpert-Raiffa (1969), instructed	23

Another experiment yielding similar results was carried out by Norman Dalkey and Bernice Brown during 1970.(6) They asked respondents for quartile estimates of uncertain quantities. Based on 1,218 cases, they found that the true answer fell outside the interquartile range 69 percent of the time (rather than the 50 percent expected from "perfect" respondents). This particular aspect of their experiments has not been published.

The only experiments that would even superficially appear to contradict the hypothesis that most people report overly "tight" subjective probability distributions for uncertain quantities, to our knowledge, were made by Ward Edwards.(7) Edwards found that, when respondents are asked to estimate, for example, the probability that a word in a given piece of prose begins with the bigram "he," and then are given additional relevant information (for example, the probability that the word begins with "h") and asked to estimate the probability again, they do not tighten up their original estimate as much as Bayes' Theorem says they should.

Edwards' results and ours are not really contradictory: we found that respondents are not conservative enough in their a priori subjective views of the world, while Edwards found that they were too conservative in evaluating additional information. Both results could follow from

the natural human tendency to accept and internalize evidence which supports one's a priori view of the world, and to filter out evidence which contradicts it. However, a full understanding of the mechanism that causes this phenomenon depends on future experiments.

V. THE SCORING SYSTEM

The scoring system used is a continuous analogue of the quadratic, or Brier, scoring system (see Ref. 8, pp. 24-26). Much emphasis has been put on the fact that this scoring system motivates the respondent to answer honestly;[9] that is, he will maximize his expected score by reporting his true subjective distribution rather than by hedging it one way or another to take advantage of the scoring system. In this experiment the scoring system used had no influence on the subjects' behavior, as they were not even informed of the scoring system we intended to use. For our purposes, then, a more important question is: "Does the continuous quadratic scoring system really measure what the decisionmakers want from a probabilistic forecast?" We believe that the answer to this question must be a qualified "yes."

The continuous quadratic scoring system may be described heuristically as follows. The range over which the forecast is being made is divided into many small intervals. For each one of these intervals the forecaster has the opportunity to place bets that the true answer will fall within that interval. The amount which he may stake is proportional to the probability reflecting the odds at which the stake is placed (for a fuller discussion, see Ref. 8, pp. 17-21, 24-25). We assume that the forecaster maximizes his subjective expected gain by taking all wagers offered at odds better than the odds he thinks are true, and rejecting all other wagers offered. The continuous quadratic score calculates, essentially, what he would win or lose by following this policy.

Now consider the position of the decisionmaker. The policies he selects are implicitly bets that this or that outcome will take place. By using a scoring system like the quadratic, he is testing his advisors by a standard highly analogous to the standards by which his decision will ultimately be tested. Of course, it is impossible to say whether the postulated distribution of available wagers corresponds closely to that actually confronting the decisionmaker; this is why only a qualified trust can be put in the continuous quadratic scoring system, or any scoring system with a similar derivation.

The continuous quadratic scoring system takes the following explicit form. If r(t)dt is the probability density function postulated by a respondent, and x is the true answer to the question, then his score for that question is given by

$$f(x) = 2r(x) - \int_D [r(t)]^2 dt \ .$$

Note that this scoring system is not "scale-invariant" in the sense that a change in units will result in a proportional change in score. In our experiments we "normalized" the scores by choosing the scale factors so that the true answer to each question would be unity. We hoped that this would make the average scores on the fourteen questions roughly comparable. The average scores realized are shown in Table 2. Note that in twelve out of fourteen cases the average score is negative. In short, if the respondents had placed many bets on the basis of their individual distributions, they would have lost money. This should come as no surprise after seeing how often the true answer lies above the 99-percentile level or below the 1-percentile level.

Table 2

AVERAGE CONTINUOUS QUADRATIC SCORE BY QUESTION

Question Number	Subject Matter	Score		
		Low	Average	High
1	Common stock	-28.01	-3.96	16.55
2	GNP	-111.13	-7.05	29.32
3	Deflated GNP	-281.69	-24.34	190.34
4	Consumer prices	-87.99	2.00	72.09
5	Unemployment	-26.93	-3.70	7.09
6	Military spending	-143.53	-13.34	12.61
7	Draft calls	-3.52	-.58	2.93
8	SVN deaths	-3.62	-.85	.22
9	Troops in SVN	-59.77	-6.40	6.98
10	Governor	-3.70	-1.34	2.81
11	Calif. Senator	-40.91	-4.85	9.94
12	Calif. Senator	-40.77	-1.82	14.45
13	Representatives	-95.63	-10.63	6.83
14	Senators	-20.25	.92	15.18

The average score of individuals varied widely; Table 3 shows that eight out of our thirty-one respondents would have been able to make money betting on their forecasts. The highest-scoring individual under the continuous quadratic scoring system (number 3 in the afternoon group) was also the highest-scoring individual on the current events quiz. However, the scores on the current events quiz were not well-correlated with the scores on the forecasting task ($r = .08$). For example, the fourth best forecaster had a score of only 15 on the current events quiz.

Table 3

AVERAGE CONTINUOUS QUADRATIC SCORE BY INDIVIDUAL

Morning Group		Afternoon Group	
Individual	Average Score	Individual	Average Score
1	-1.30	1	-3.25
2	6.82	2	-.78
3	-3.45	3	15.87
4	-10.91	4	-19.07
5	4.64	5	-8.99
6	-4.53	6	-20.41
7	-13.85	7	-7.76
8	-1.72	8	-21.18
9	-13.09	9	-27.45
10	-17.00	10	-3.15
11	1.93	11	.97
12	-9.29	12	-4.77
13	-8.34	13	.84
14	-4.89	14	-.54
15	-1.64	15	8.57
		16	3.68

VI. CONSENSUS FORECASTS

To what extent could more accurate, higher-scoring forecasts be achieved by combining the forecasts of many individuals into a single "consensus forecast?" There are, of course, many techniques by which this consensus could be effected. One good way is to average the probability densities inferred from each individual forecast. The payoff to this forecast will always be greater than the average payoff to individual forecasters. To see this, assume there are N forecasters, and the *ith* forecaster gives a density function $r_i(t)$. The average density will be

$$r(t) = \frac{1}{N}\sum_{i=1}^{n} r_i(t) \ .$$

Thus we see that

(payoff to consensus) - (average payoff)

$$= 2r(x) - \int_D [r(t)]^2 dt - \frac{2}{N}\sum_{i=1}^{n} r_i(x) + \int_D \sum_{i=1}^{n} \frac{\left[r_i(t)\right]^2}{N}\, dt$$

$$= \int_D \sum_{i=1}^{n} \frac{\left[r_i(t)\right]^2}{N} - \left[\frac{\sum_{i=1}^{n} r_i(t)}{N}\right]^2 dt$$

$$= \frac{1}{N}\int_D \sum_{i=1}^{n} (r_i(t) - r(t))^2 \, dt \geq 0 \ .$$

It is an amusing fact that we can calculate the amount of improvement which the consensus distribution's score will show over the average of the individuals' scores without knowing what the true answer may be!

There is a practical drawback to using the average probability density as a consensus device: when the initial responses are in the form of percentile breaks, it takes a lot of calculation to convert these to density functions and then average the results. A more straightforward device is to simply average the percentile breaks. For example, if a(.01,i) represents the .01 percentile on a question reported by the *ith* respondent, then the .01 percentile on the consensus distribution would be

$$\sum_{i=1}^{n} \frac{a(.01,i)}{n} .$$

Although this method is computationally convenient, at first glance it appears that it would not generate a consensus distribution with as much "spread" as that generated by the average density method. Since the individual forecasters have a tendency not to spread their distributions adequately, one might think that this would be an undesirable tendency. Table 4 shows that the average density method does outscore the average percentile break method in most cases, and that both of them outscore the average individual by a wide margin.

This result is similar to those in most Delphi exercises associated with point estimates. The group consensus, however constructed, will outscore 80 percent to 90 percent of the group.

It is interesting to compare Table 4 with Table 3. Four individuals out of thirty-one were able to outscore both consensus distributions (on the average). These four individuals seem to have no obvious distinguishing characteristic in common, as can be seen from the following table.

Sex	Age	Academic Status	Current Events Quiz Score	Major
F	41	Graduate	29	Counselor Ed.
F	23	Graduate	15	Opera
M	21	Senior	37	History
M	20	Senior	18	Economics

Table 4

COMPARATIVE SCORES OF TWO CONSENSUS METHODS

Question	Consensus Methods		Individual's Average
	Average Density	Average Breaks	
1 Common stock	.53	.25	-3.96
2 GNP	10.09	6.20	-7.05
3 Deflated GNP	22.40	9.11	-24.34
4 Consumer prices	20.39	12.21	2.00
5 Unemployment	-.07	.13	-3.70
6 Military spending	1.00	10.77	-13.34
7 Draft calls	.22	-.66	-.58
8 SVN deaths	-.10	-.32	-.85
9 Troops in SVN	.80	2.40	-6.40
10 Governor	-.04	-.19	-1.34
11 Calif. Senator	.73	1.22	-4.85
12 Calif. Governor	3.00	2.38	-1.82
13 Representatives	-.39	4.20	-10.63
14 Senators	3.72	5.36	.92
Total	62.28	53.06	-75.94
Average (Total ÷ 14)	4.45	3.79	-5.42

The reader may ask whether any gain in accuracy has been achieved by asking the respondents for a probability distribution as opposed to simple point estimates. Even if you feel (as we do) that a probability distribution is a much more meaningful and useful forecast than a point estimate, still it is possible to construct such a distribution from point estimates if you have enough of them. Could such an approach score as well or better than our distributional approach? To test this idea we imagined that each respondent had given as a point estimate on each question the median (50-percentile estimate) of his distribution for that question. If you imagine that these estimates were drawn from a normal distribution, then the best estimate of the parameters (mean and variance) of that distribution would be

$$m = \bar{x} = \frac{\sum x_i}{N}$$

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{N-1} .$$

How would the normal distribution with these parameters score under the continuous version of the quadratic scoring system? Table 5 compares the score achieved by this method with the two other consensus methods we have discussed. The scores for this method appear under the heading "total normal." Note that, on the average, this method scores much lower than the "average density" method, and somewhat lower than the "average breaks" method.

Table 5

COMPARATIVE SCORES OF FOUR CONSENSUS METHODS

Question	Consensus Methods			
	Using Whole Distribution		Using Medians Only	
	Average Density	Average Breaks	Total Normal	Average of Normals
1 Common stock	.53	.25	.55	1.58
2 GNP	10.09	6.20	2.78	2.75
3 Deflated GNP	22.40	9.11	13.06	12.90
4 Consumer prices	20.39	12.21	11.27	10.48
5 Unemployment	-.07	.13	-.95	.45
6 Military spending	1.00	10.77	7.91	5.70
7 Draft calls	.22	-.66	.21	.28
8 SVN deaths	-.10	-.32	-.01	.09
9 Troops in SVN	.80	2.40	2.21	1.55
10 Governors	-.04	-.19	-.11	.31
11 Calif. Senator	.73	1.22	.27	1.64
12 Calif. Governor	3.00	2.38	6.13	4.07
13 Representatives	-.39	4.20	2.89	4.16
14 Senators	3.72	5.36	3.09	2.65
Total	62.28	53.06	49.30	48.61
Average (Total ÷ 14)	4.45	3.79	3.52	3.47

The superior performance of the "average density" method leads us to try the following expedient: presume that each individual gave a normal distribution, with mean equal to his 50-percentile estimate, and standard deviation equal to that calculated above. Then take the average of the corresponding probability density functions as the probability density function of your consensus forecast. It is clear that this formula will spread the probability density out more than any other consensus method we have considered. The score achieved by this "average of normals" consensus is shown in column 4 of Table 5. Its average is the worst of the four methods we have considered, but on the other hand it is the only method which achieves a positive score on all fourteen questions.

In summary, the best of the consensus methods we have considered from the standpoint of average score is the "average density" method, using the entire distributions reported by respondents. The best from the standpoint of maximizing the minimum score is the very conservative "average of normals" method.

Appendix A

MAIN QUESTIONNAIRE

(Sample Question)

0. PRODUCTION OF ANTHRACITE COAL IN THE U.S.

(Millions of tons)

1945	54.4
1950	44.1
1955	26.2
1960	18.8
1965	14.9

How much anthracite coal will be produced in the U.S. during 1970?

1% chance of being less than _____.
10% " " " " " _____.
30% " " " " " _____.
50% " " " " " _____.
70% " " " " " _____.
90% " " " " " _____.
99% " " " " " _____.

How well acquainted are you with the subject matter of this question?

_____

1. COMMON STOCK PRICES

(Standard & Poor's 500-stock average)

	March	June	September	December
1962	70.29	55.63	58.00	62.64
1963	65.67	70.11	72.85	74.17
1964	78.80	80.24	83.41	83.96
1965	86.83	85.04	89.38	91.73
1966	88.88	86.06	77.81	81.33
1967	89.42	91.43	95.81	95.30
1968	89.09	100.53	101.34	106.48
1969	99.30	99.14	94.51	91.11
1970	88.65	76.40		

What will be the value of Standard and Poor's 500 stock average in December 1970?

1% chance of being less than _____.
10% " " " " " _____.
30% " " " " " _____.
50% " " " " " _____.
70% " " " " " _____.
90% " " " " " _____.
99% " " " " " _____.

How well acquainted are you with the subject matter of this question?

_____.

True Answer = 90.05

2. U.S. GROSS NATIONAL PRODUCT

(Billions of dollars, annual rate)

1965	684.9
1966	749.9
1967	793.5
1968	865.7
First Quarter, 1969	908.7
Second Quarter, 1969	924.8
Third Quarter, 1969	942.8
Fourth Quarter, 1969	952.2
First Quarter, 1970	960.4

What will be the annual rate of the U.S. Gross National Product during the fourth quarter, 1970?

1% chance of being less than _____.
10% " " " " " _____.
30% " " " " " _____.
50% " " " " " _____.
70% " " " " " _____.
90% " " " " " _____.
99% " " " " " _____.

How well acquainted are you with the subject matter of this question?

_____

True Answer = 990.9

3. U.S. GROSS NATIONAL PRODUCT IN 1958 DOLLARS

(Billions of 1958 dollars, annual rate)

1965	617.8
1966	658.1
1967	674.6
1968	707.6
First Quarter, 1969	723.1
Second Quarter, 1969	726.7
Third Quarter, 1969	730.6
Fourth Quarter, 1969	729.8
First Quarter, 1970	726.9

What will be the annual rate of the U.S. Gross National Product in 1958 dollars during the fourth quarter of 1970?

1% chance of being less than _____.
10% " " " " " _____.
30% " " " " " _____.
50% " " " " " _____.
70% " " " " " _____.
90% " " " " " _____.
99% " " " " " _____.

How well acquainted are you with the subject matter of this question?

_____

True Answer = 721.3

4. CONSUMER PRICE INDEX (all items: 1957-59 = 100)

	March	June	September	December
1962	105.0	105.3	106.1	105.8
1963	106.2	106.6	107.1	107.6
1964	107.7	108.0	108.4	108.8
1965	109.0	110.1	110.2	111.0
1966	112.0	112.9	114.1	114.7
1967	115.0	116.0	117.1	118.2
1968	119.5	120.9	122.2	123.7
1969	125.6	127.6	129.3	131.3
1970	133.2	135.2		

Where will the consumer price index stand in December 1970?

1% chance of being less than _____.
10% " " " " " _____.
30% " " " " " _____.
50% " " " " " _____.
70% " " " " " _____.
90% " " " " " _____.
99% " " " " " _____.

How well acquainted are you with the subject matter of this question?

_____

True Answer = 138.9

5. U.S. UNEMPLOYMENT (percent of civilian labor force)

(Not seasonally adjusted)

	March	June	September	December
1960	6.2	6.1	4.8	6.4
1961	7.7	7.5	5.7	5.8
1962	6.2	6.0	4.9	5.3
1963	6.3	6.4	4.8	5.3
1964	5.9	6.1	4.5	4.7
1965	5.1	5.5	3.8	3.8
1966	4.0	4.5	3.3	3.5
1967	3.9	4.6	3.7	3.5
1968	3.8	4.5	3.3	3.1
1969	3.5	4.1	3.7	3.2
1970	4.6			

What will the unemployment rate be in December, 1970?

1% chance of being less than _____.
10% " " " " " _____.
30% " " " " " _____.
50% " " " " " _____.
70% " " " " " _____.
90% " " " " " _____.
99% " " " " " _____.

How well acquainted are you with the subject matter of this question?

_____

True Answer = 5.6

6. U.S. MILITARY SPENDING

(Billions of dollars, annual rate)

1965	50.1
1966	60.7
1967	72.4
1968	78.0
First Quarter, 1969	79.0
Second Quarter, 1969	78.5
Third Quarter, 1969	80.3
Fourth Quarter, 1969	79.2
First Quarter, 1970	77.3

What will the annual rate of U.S. military spending be in the fourth quarter of 1970?

1% chance of being less than _____.
10% " " " " " _____.
30% " " " " " _____.
50% " " " " " _____.
70% " " " " " _____.
90% " " " " " _____.
99% " " " " " _____.

How well acquainted are you with the subject matter of this question?

_____

True Answer = 74.6

7. SELECTIVE SERVICE CALLS

(Total calls, by month, in thousands)

	1965	1966	1967	1968	1969	1970
January	5.4	29.3	15.6	34.0	26.8	12.5
February	3.0	22.4	10.9	23.3	33.7	19.0
March	7.9	18.4	11.9	41.0	33.1	19.0
April	13.7	19.2	11.4	44.0	33.0	19.0
May	15.1	40.6	18.0	44.0	27.6	15.0
June	17.0	18.5	19.8	20.0	25.9	15.0
July	17.1	28.5	19.9	15.0	22.3	15.0
August	16.5	36.6	29.0	18.3	29.5	
September	27.4	37.3	25.0	12.2	29.0	
October	29.0	49.2	17.0	13.8	10.0	
November	34.3	37.6	22.0	10.0	10.0	
December	40.2	12.1	18.2	15.0	9.0	

What will the total Selective Service call be for December, 1970?

1% chance of being less than _____.
10% " " " " " _____.
30% " " " " " _____.
50% " " " " " _____.
70% " " " " " _____.
90% " " " " " _____.
99% " " " " " _____.

How well acquainted are you with the subject matter of this question?

_____

True Answer = 7.0

8. U.S. MILITARY DEATHS IN SOUTHEAST ASIA FROM HOSTILE CAUSES

	Monthly Average
1961-63	3
1964	12
1965	114
1966	417
1967	781
1968	1216
1969	784
1970 (first six months)	497

How many U.S. military deaths in Southeast Asia from hostile causes will occur in December 1970?

1% chance of less than _____.
10% " " " " _____.
30% " " " " _____.
50% " " " " _____.
70% " " " " _____.
90% " " " " _____.
99% " " " " _____.

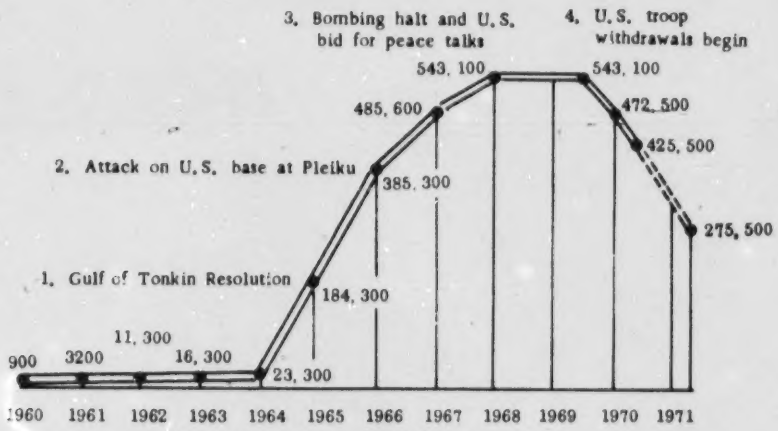
How well acquainted are you with the subject matter of this question?

_____

True Answer = 153

9. U.S. DEPLOYMENTS IN VIETNAM

U.S. Troops in Vietnam: Escalation and Withdrawal *



1. After a reported attack on U.S. destroyers in the Gulf of Tonkin by North Vietnamese torpedo boats, Congress authorized President Johnson to "repel any armed attack against the forces of the United States and to prevent further aggression."
2. A Vietcong attack on an American Base at Pleiku was called by the U.S. a "test of will." Air attacks on North Vietnam and the dispatch of more American troops followed.
3. With protests against the war mounting, President Johnson announced a halt to bombing of North Vietnam and a new bid for peace talks.
4. At a meeting last summer with President Thieu, President Nixon announced the first withdrawals of American troops--25,000--since the Vietnam build-up began.

*New York Times, April 26, 1970.

How many U.S. military personnel will be in Vietnam in December, 1970?

1% chance of less than _____.
10% " " " " _____.
30% " " " " _____.
50% " " " " _____.
70% " " " " _____.
90% " " " " _____.
99% " " " " _____.

How well acquainted are you with the subject matter of this question?

_____

True Answer = 339,200

10. GOVERNORS OF STATES

	Democrat	Republican
1960	34	16
1962	34	16
1964	33	17
1966	25	25
1968	19	31
Today	18	32

In November, 1970, 35 governors will be elected. Eleven of these posts are now held by Democrats and 24 by Republicans.

How many Republican governors will there be in the U.S. after the election of November, 1970?

1% chance of fewer than _____.
10% " " " " _____.
30% " " " " _____.
50% " " " " _____.
70% " " " " _____.
90% " " " " _____.
99% " " " " _____.

How well acquainted are you with the subject matter of this question?

_____

True Answer = 21

11. U.S. SENATOR FROM CALIFORNIA

Year	Rep. Candidate	% of Vote	Dem. Candidate	% of Vote
1962	Kuchel	56.4	Richards	43.6
1964	Murphy	51.5	Salinger	48.5
1968	Rafferty	46.7	Cranston	52.0

What percentage of the popular vote will the Republican candidate for U.S. Senator from California receive in the November, 1970 elections?

1% chance of less than _____.
10% " " " " _____.
30% " " " " _____.
50% " " " " _____.
70% " " " " _____.
90% " " " " _____.
99% " " " " _____.

How well acquainted are you with the subject matter of this question?

_____

True Answer = 44.4

12. PERCENT OF VOTE CAST FOR REPUBLICAN CANDIDATE FOR GOVERNOR OF CALIFORNIA

Year	Republican Candidate	% of Vote Received
1954	Goodwin Knight	56.4
1958	William Knowland	41.1
1962	Richard Nixon	47.7
1966	Ronald Reagan	57.7

What percentage of the popular vote will the Republican candidate for Governor of California get in the election of November, 1970?

1% chance of receiving less than _____%.
10% " " " " " _____%.
30% " " " " " _____%.
50% " " " " " _____%.
70% " " " " " _____%.
90% " " " " " _____%.
99% " " " " " _____%.

How well acquainted are you with the subject matter of this question?

_____

True Answer = 52.9

13. HOUSE OF REPRESENTATIVES

After Election of	Dem.	Rep.	Misc.	After Election of	Dem.	Rep.
1950	234	199	2	1960	263	174
1952	213	221	1	1962	259	176
1954	232	203		1964	295	140
1956	234	201		1966	248	187
1958	283	154		1968	243	192

How many Democrats will there be in the House of Representatives after the election of November, 1970?

1% chance of less than _____.
10% " " " " _____.
30% " " " " _____.
50% " " " " _____.
70% " " " " _____.
90% " " " " _____.
99% " " " " _____.

How well acquainted are you with the subject matter of this question?

_____

True Answer = 255

14. U.S. SENATE

After Election of	Dem.	Rep.	Misc.	After Election of	Dem.	Rep.
1950	48	47	1	1960	64	36
1952	47	48	1	1962	68	32
1954	48	47	1	1964	68	32
1956	49	47		1966	64	36
1958	66	34		1968	58	42

There are 35 seats up for election in 1970. Of these, 25 are now held by Democrats and 10 by Republicans.

How many Democrats will there be in the Senate after the election of 1970?

1% chance of fewer than _____.
10% " " " " _____.
30% " " " " _____.
50% " " " " _____.
70% " " " " _____.
90% " " " " _____.
99% " " " " _____.

How well acquainted are you with the subject matter of this question?

_____

True Answer = 54

15. PEACE TALKS

What will the status of the Paris Peace Talks be on December 31, 1970?

They will have been called off in anger by the North Vietnam and National Liberation Front side.	_____%
They will have been called off in anger by the South Vietnam and U.S. side.	_____%
They will have been adjourned by mutual agreement, but with no peace treaty or armistice signed.	_____%
They will have been adjourned, a peace treaty or armistice having been signed.	_____%
They will be continuing, a cease-fire having been signed.	_____%
They will be continuing, but show little more signs of success than they do now.	_____%
Some status not covered by the above will exist.	_____%
	Sum: 100

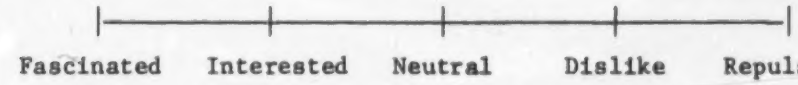
How well acquainted are you with the subject matter of this question?

_____

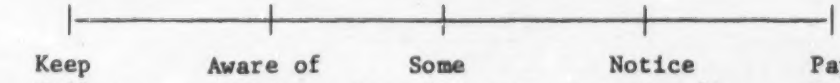
True Answer = Continuing, with little sign of success

ATTITUDE AND ACQUAINTANCE SCALES

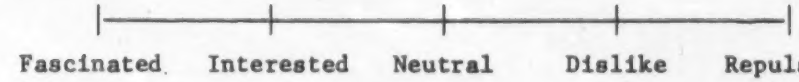
Attitude toward Economic Affairs:



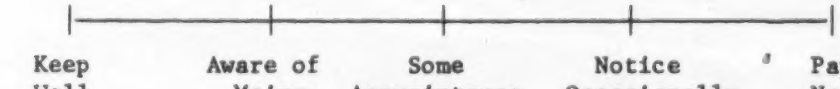
Acquaintance with Economic Affairs:



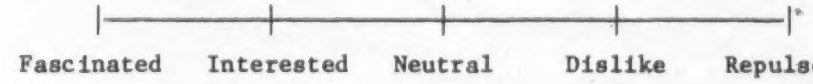
Attitude toward Domestic Politics:



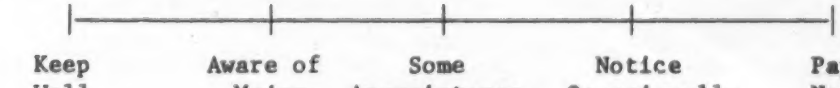
Acquaintance with Domestic Politics:



Attitude toward Foreign Affairs:



Acquaintance with Foreign Affairs:



Appendix B

CURRENT EVENTS BACKGROUND QUIZ

INSTRUCTIONS

Current Events Background Quiz

Please black in the letter on the answer sheet corresponding to the best answer. For example:

0. Who is Abraham Lincoln?

(a) Republican president
(b) Spanish leader of 1930's
(c) A New Jersey Nazi
(d) Henry Ford aide

0. (●) (b) (c) (d)

Answer all questions, even if it's only a guess.

Current Events Background Quiz

1. Who is John Mitchell?

   (a) Attorney General
   (b) Film star
   (c) Martyred bomber advocate
   (d) Popular writer of 40's and 50's

2. What is NATO?

   (a) Leftist information agency
   (b) Navy quick take-off gadget
   (c) International group of central banks
   (d) Organization of some western nations

3. Who is David Kennedy?

   (a) Chairman, Chase-Manhattan Bank
   (b) Secretary of the Treasury
   (c) Republican uncle of Ted Kennedy
   (d) Chairman, CEA

4. Who is Mike Mansfield?

   (a) Governor of Montana
   (b) Labor leader
   (c) Columnist for the Washington Post
   (d) Senate leader

5. Who is Golda Meir?

   (a) Israeli Premier
   (b) UN peacemaker
   (c) Assistant Secretary of State
   (d) German martyr

6. In what year did the greatest one day increase in the Dow-Jones Industrial Average occur?

   (a) 1970
   (b) 1929
   (c) 1963
   (d) 1967

7. Who is Benjamin Aaron?

   (a) Gaullist journalist
   (b) Indicted financier
   (c) Big league outfielder
   (d) Teachers' strike mediator

8. Who is Henry Kissinger?

   (a) ABM foe
   (b) Presidential aide
   (c) Assistant Secretary of Defense
   (d) West German Chancellor

9. The best-known ratings of corporate and municipal bonds are put out by

   (a) Dow-Jones
   (b) Dun & Bradstreet
   (c) Standard & Poor
   (d) Moody

10. Who is Walter Hickel?

   (a) Secretary of Labor
   (b) Secretary of Agriculture
   (c) Former Alaskan Governor
   (d) Federal Youth Advisor

11. Who is Bertrand Russell?

   (a) Inventor
   (b) Defense lawyer
   (c) Philosopher
   (d) Foreign Minister under Asquith

12. Deflation is usually accompanied by:

   (a) Increased confidence
   (b) More pay for service workers
   (c) Increased imports
   (d) High unemployment

13. Who is Winthrop Rockefeller?

   (a) Arkansas governor
   (b) Old-time oil baron
   (c) New York banker
   (d) Youthful legislator

14. Who is E. G. Marshall?

   (a) Chief of Staff in World War II
   (b) TV star
   (c) Secretary of State in late 40's
   (d) Liberal attorney

15. What is the current status of rent control?

(a) It is still a law in California, but not enforced by the present State government.
(b) The courts have declared it unconstitutional except in time of war.
(c) It exists in some cities, but not in most.
(d) The president can impose it at any time.

16. Who is now governor of Ohio?

(a) Rhodes
(b) Ogilvie
(c) Sargeant
(d) Young

17. Who is Norodom Sihanouk?

(a) Philippine pacifist philosopher
(b) Ousted Cambodian Prince
(c) Imprisoned Thai politician
(d) Laotian neutralist leader

18. The items used in calculating BLS consumer price index:

(a) Have not changed since 1958
(b) Vary from time to time to reflect changes in consumer spending patterns
(c) Do not include services such as medical care
(d) Include only necessities

19. How many members are there in the California State Senate?

(a) 30
(b) 40
(c) 93
(d) 120

20. Who is Allen Dulles?

(a) Former Undersecretary of State
(b) Famous judge
(c) Master spy
(d) Pentagon leader

21. What is the effect on bonds when investors expect inflation?

(a) Bond prices and yields both go down.
(b) Bond prices go up and yields go down.
(c) Bond prices go down and yields go up.
(d) Bond prices and yields both go up.

22. What was Proposition 7 in the June election?

(a) To eliminate archaic and redundant provisions in the State Constitution
(b) To increase state aid for schools and welfare
(c) To limit off-shore oil drilling
(d) To raise interest ceiling on state bonds

23. What is the "Nixon Doctrine"?

(a) Use any degree of force necessary to preserve orderly academic processes
(b) Attack enemy sanctuaries whenever doing so will forestall attacks on American forces
(c) Consult military advisors on military problems
(d) Provide aid, but not ground forces, to threatened friendly nations

24. What does "short selling" mean?

(a) Selling early to avoid capital gains taxes
(b) Selling borrowed stock
(c) Forced selling of margined stock
(d) Selling hastily in a declining market

25. Who is Mark Hanna?

(a) Prize-winning researcher
(b) Embattled college president
(c) Old-time Republican
(d) Youth leader

26. What kind of jet bombers are used by Israel?

(a) Mystique III-c
(b) Il-28
(c) B-58
(d) F-4

27. Who is Milton Friedman?

(a) Accused ax-murderer
(b) Comic adman
(c) New left poet
(d) Conservative economist

28. Who is Robert Monagan?

(a) Assembly leader
(b) Democratic legislator
(c) Former Secretary of Commerce
(d) White House domestic aide

29. Who is Tran Ngoc Chau?

(a) Non-Communist foe of Thieu
(b) Political Chief of COSVN
(c) Would-be assassin of Marshall Ky
(d) Chief of Saigon Secret Police

30. Who is Paul W. McCracken?

(a) Governor of Indiana
(b) Chairman of the Council of Economic Advisors
(c) Secretary of Commerce
(d) Senator from Idaho

31. Who is Larry O'Brien?

(a) Democratic National Chairman
(b) Nixon Aide
(c) Mayor of Boston
(d) Accused racketeer

32. Who is J. J. Sisco?

(a) Old-time Mexican bandit
(b) Crime fighter
(c) Assistant Secretary of State
(d) Conglomerate titan

33. What was the Gross National Product in 1969?

(a) 120.8 billion dollars
(b) A little over 900 billion dollars
(c) A little over 200 billion dollars
(d) More than a trillion dollars

34. Who is Ed Reinecke?

(a) Chairman of Republican National Committee
(b) Defeated candidate for Attorney General
(c) Congressman from Northern California
(d) Lt. Governor of California

35. Who is Kim Il-Sung?

(a) Neutralist strongman
(b) North Korean leader
(c) Imprisoned terrorist
(d) Electronic musician

36. How old do you have to be in order to be classed as unemployed by the Bureau of Labor Statistics?

(a) 16
(b) 21
(c) 18
(d) 14

37. What is the "Eleventh Commandment"?

(a) "Never volunteer."
(b) "Don't speak ill of any fellow Republican."
(c) "Get the money."
(d) "Cross no lawful picket lines."

38. What is an M-48?

(a) North Vietnamese rocket launcher
(b) U.S. tank
(c) British jet engine
(d) Israeli sub-machine gun

39. Who is Bernard Cornfeld?

(a) Las Vegas gambler
(b) Nixon confidante
(c) Alleged atom-spy
(d) Ousted IOS leader

40. Who suggested "benign neglect" of racial problems?

(a) Murphy
(b) Douglas
(c) Moynihan
(d) Wallace

41. What does "AK-47" mean?

(a) An Israeli armored vehicle
(b) A rocket used by Viet Cong
(c) A type of rifle
(d) A Chinese land mine

42. Who is J. M. Keynes?

(a) British economist
(b) Harvard professor
(c) New Deal cabinet member
(d) French labor leader

43. Which of the following appears to be carrying out the political theories of Kevin Phillips?

(a) Agnew
(b) Ted Kennedy
(c) Rubin
(d) Moynahan

44. What does MRN stand for?

(a) Main Reactor Nucleus
(b) National Republican Movement
(c) Moratorium Radio Network
(d) Make Revolution Now

45. Which of the follcwing would not be classified by the Bureau of Labor Statistics as "unemployed"?

(a) A man looking for a job, who now works half-time in his dad's gas station without pay.
(b) A man not working, but waiting to report to a new job scheduled to start in two weeks.
(c) A man not working, but waiting to be called back to a job from which he has been laid off.
(d) A married woman whose husband is employed, who is herself not working but seeking part-time work as a clerk.

Answer Sheet

1. (a) (b) (c) (d)	16. (a) (b) (c) (d)	31. (a) (b) (c) (d)
2. (a) (b) (c) (d)	17. (a) (b) (c) (d)	32. (a) (b) (c) (d)
3. (a) (b) (c) (d)	18. (a) (b) (c) (d)	33. (a) (b) (c) (d)
4. (a) (b) (c) (d)	19. (a) (b) (c) (d)	34. (a) (b) (c) (d)
5. (a) (b) (c) (d)	20. (a) (b) (c) (d)	35. (a) (b) (c) (d)
6. (a) (b) (c) (d)	21. (a) (b) (c) (d)	36. (a) (b) (c) (d)
7. (a) (b) (c) (d)	22. (a) (b) (c) (d)	37. (a) (b) (c) (d)
8. (a) (b) (c) (d)	23. (a) (b) (c) (d)	38. (a) (b) (c) (d)
9. (a) (b) (c) (d)	24. (a) (b) (c) (d)	39. (a) (b) (c) (d)
10. (a) (b) (c) (d)	25. (a) (b) (c) (d)	40. (a) (b) (c) (d)
11. (a) (b) (c) (d)	26. (a) (b) (c) (d)	41. (a) (b) (c) (d)
12. (a) (b) (c) (d)	27. (a) (b) (c) (d)	42. (a) (b) (c) (d)
13. (a) (b) (c) (d)	28. (a) (b) (c) (d)	43. (a) (b) (c) (d)
14. (a) (b) (c) (d)	29. (a) (b) (c) (d)	44. (a) (b) (c) (d)
15. (a) (b) (c) (d)	30. (a) (b) (c) (d)	45. (a) (b) (c) (d)

REFERENCES

1. Derouin, R. G., and G. F. Cobb, *Weather Bureau Forecast Verification Scores 1968-69 and Some Performance Trends from 1966*, ESSA WBTM FCST 15, U.S. Department of Commerce, May 1970.

2. Winkler, R. L., and A. H. Murphy, "Evaluation of Subjective Precipitation Probability Forecasts," *Proceedings of the First National Conference on Statistical Meteorology*, American Meteorological Society, Boston, 1968, pp. 148-157.

3. Slovic, P., and S. Lichtenstein, "Importance of Variance Preferences in Gambling Decision," *Journal of Experimental Psychology*, 1968, Vol. 78, No. 4, pp. 646-654.

4. Slovic, P., and S. Lichtenstein, "Relative Importance of Probabilities and Payoffs in Risk Taking," *Journal of Experimental Psychology Monograph*, Vol. 78, No. 3, Part 2, November 1968.

5. Alpert, M., and H. Raiffa, "A Progress Report on the Training of Probability Assessors," unpublished note, August 28, 1969.

6. Dalkey, N., and B. Brown, *Comparison of Group Judgment Techniques with Short-Range Predictions and Almanac Questions*, The Rand Corporation, R-678-ARPA, January 1971.

7. Edwards, W., *Nonconservative Probabilistic Information Processing Systems*, Decision Sciences Laboratory, Electronic Systems Division, Air Force Systems Command, ESD-TR-66-404, December 1966.

8. Brown, T. A., *Probabilistic Forecasts and Reproducing Scoring Systems*, The Rand Corporation, RM-6299-ARPA, June 1970.

9. Shuford, E. H., A. Albert, and H. E. Massengill, "Admissible Probability Measurement Procedures," *Psychometrika*, Vol. 31, No. 2, June 1966.